



1638

1600

ENTERED

DATE: 03/13/2002
TIME: 14:59:12R.S.
RECEIVED

MAR 22 2002

TECH CENTER 1600/2900

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/657,631Input Set : A:\-798-2.app
Output Set: N:\CRF3\03132002\I657631.raw

3 <110> APPLICANT: Etzler, Marilynn E.
 4 Roberts, Nicholas J.
 5 The Regents of the University of California
 7 <120> TITLE OF INVENTION: LNP, a Protein Involved in the Initiation of
 8 Mycorrhizal Infection in Plants
 10 <130> FILE REFERENCE: 023070-079820US
 12 <140> CURRENT APPLICATION NUMBER: US 09/657,631
 13 <141> CURRENT FILING DATE: 2000-09-06
 15 <160> NUMBER OF SEQ ID NOS: 14
 17 <170> SOFTWARE: PatentIn Ver. 2.1
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 1643
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Dolichos biflorus
 24 <220> FEATURE:
 25 <221> NAME/KEY: CDS
 26 <222> LOCATION: (51)..(1439)
 27 <223> OTHER INFORMATION: lectin nucleotide phosphohydrolase (LNP, NBP46 or
 28 DB46) root lectin
 30 <220> FEATURE:
 31 <221> NAME/KEY: mat_peptide
 32 <222> LOCATION: (195)..(1436)
 34 <400> SEQUENCE: 1
 35 gaaactgaaa cgagtactct ttcagtgggt aggttctgag agattcagaa atg aat 56
 36 Met Asn
 38 tgg gtg tgg cca aag aca aag agc atg agc ttc cta ctc ctc atc act 104
 39 Trp Val Trp Pro Lys Thr Lys Ser Met Ser Phe Leu Leu Leu Ile Thr
 40 -45 -40 -35
 42 ttt cta ctc ttc tca ttg cca aaa ctt tct tct tcg caa tat gtt ggg 152
 43 Phe Leu Leu Phe Ser Leu Pro Lys Leu Ser Ser Ser Gln Tyr Val Gly
 44 -30 -25 -20 -15
 46 aac agt atc tta cta aat cat cgt aag ata ctt ccc aac cag gaa ctc 200
 47 Asn Ser Ile Leu Leu Asn His Arg Lys Ile Leu Pro Asn Gln Glu Leu
 48 -10 -5 -1 1
 50 ctt acc tct tac gct gtc atc ttt gat gct ggt agc tct ggg agt cgt 248
 51 Leu Thr Ser Tyr Ala Val Ile Phe Asp Ala Gly Ser Ser Gly Ser Arg
 52 5 10 15
 54 gtc cat gtc ttc aat ttt gac cag aac tta gat ctc ctg cac att ggc 296
 55 Val His Val Phe Asn Phe Asp Gln Asn Leu Asp Leu Leu His Ile Gly
 56 20 25 30
 58 aat gac ctc gag ttt aca aaa aag atc aaa ccc ggt ttg agc tca tac 344
 59 Asn Asp Leu Glu Phe Thr Lys Lys Ile Lys Pro Gly Leu Ser Ser Tyr
 60 35 40 45 50

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/657,631

DATE: 03/13/2002
TIME: 14:59:12

Input Set : A:\-798-2.app
Output Set: N:\CRF3\03132002\I657631.raw

62	gct	gat	aag	cct	gaa	aaa	gct	gca	gaa	tct	ctc	att	cca	ctt	ttg	gag	392
63	Ala	Asp	Lys	Pro	Glu	Lys	Ala	Ala	Glu	Ser	Leu	Ile	Pro	Leu	Leu	Glu	
64					55				60						65		
66	gaa	gct	gaa	gat	gtt	gtc	cct	gag	gaa	ctg	cac	ccc	aag	aca	ccc	ctt	440
67	Glu	Ala	Glu	Asp	Val	Val	Pro	Glu	Glu	Leu	His	Pro	Lys	Thr	Pro	Leu	
68					70				75						80		
70	aag	ctt	ggg	gca	aca	gca	ggt	ttg	agg	ctc	ttg	gat	ggg	gat	gct	gct	488
71	Lys	Leu	Gly	Ala	Thr	Ala	Gly	Leu	Arg	Leu	Leu	Asp	Gly	Asp	Ala	Ala	
72					85				90						95		
74	gaa	aag	ata	ttg	caa	gca	gct	gtt	agg	gaa	atg	ttc	agg	aac	aga	agt	536
75	Glu	Lys	Ile	Leu	Gln	Ala	Val	Arg	Glu	Met	Phe	Arg	Asn	Arg	Ser	Ser	
76					100				105						110		
78	ctg	agc	gtt	caa	cct	gat	gca	gta	tct	gtt	att	gat	gga	acc	caa	gaa	584
79	Leu	Ser	Val	Gln	Pro	Asp	Ala	Val	Ser	Val	Ile	Asp	Gly	Thr	Gln	Glu	
80	115				120					125					130		
82	ggt	tct	tac	tta	tgg	gtt	aca	gtt	aac	tat	ctg	tta	gga	aag	ttg	gga	632
83	Gly	Ser	Tyr	Leu	Trp	Val	Thr	Val	Asn	Tyr	Leu	Leu	Gly	Lys	Leu	Gly	
84					135					140					145		
86	aag	aag	ttt	aca	aaa	act	gtg	gga	gtg	ata	gat	ctt	gga	ggt	gct	tca	680
87	Lys	Lys	Phe	Thr	Lys	Thr	Val	Gly	Val	Ile	Asp	Leu	Gly	Gly	Ala	Ser	
88					150					155					160		
90	gtt	caa	atg	gct	tat	gct	gtc	tca	aga	aat	aca	gct	aaa	aat	gcc	cca	728
91	Val	Gln	Met	Ala	Tyr	Ala	Val	Ser	Arg	Asn	Thr	Ala	Lys	Asn	Ala	Pro	
92					165				170						175		
94	aaa	cca	cca	caa	gga	gag	gat	cca	tac	atg	aag	aag	ctt	gtc	ctc	aag	776
95	Lys	Pro	Pro	Gln	Gly	Glu	Asp	Pro	Tyr	Met	Lys	Lys	Leu	Val	Leu	Lys	
96					180				185						190		
98	gga	aag	aaa	tat	gac	ctt	tat	gtt	cac	agt	tac	ttg	cgt	tat	ggt	aac	824
99	Gly	Lys	Lys	Tyr	Asp	Leu	Tyr	Val	His	Ser	Tyr	Leu	Arg	Tyr	Gly	Asn	
100	195				200					205					210		
102	gac	gca	gca	cgt	gtt	aag	att	ttt	aag	acc	act	gat	ggt	gct	gct	agt	872
103	Asp	Ala	Ala	Arg	Val	Lys	Ile	Phe	Lys	Thr	Thr	Asp	Gly	Ala	Ala	Ser	
104					215					220					225		
106	cct	tgt	cta	ttg	gca	ggc	tat	gaa	gat	ata	tac	aga	tat	tcc	gga	gaa	920
107	Pro	Cys	Leu	Leu	Ala	Gly	Tyr	Glu	Asp	Ile	Tyr	Arg	Tyr	Ser	Gly	Glu	
108					230					235					240		
110	tcg	tac	aat	atc	tat	ggt	ccc	act	tct	ggt	gcc	aac	ttt	aat	gag	tgc	968
111	Ser	Tyr	Asn	Ile	Tyr	Gly	Pro	Thr	Ser	Gly	Ala	Asn	Phe	Asn	Glu	Cys	
112					245				250						255		
114	cgt	gac	cta	gct	ctt	cag	att	ctc	aga	ttg	aat	gag	cca	tgt	tcc	cat	1016
115	Arg	Asp	Leu	Ala	Leu	Gln	Ile	Leu	Arg	Leu	Asn	Glu	Pro	Cys	Ser	His	
116					260				265						270		
118	gaa	aac	tgc	acc	ttt	ggt	ggg	ata	tgg	gat	ggt	gga	aaa	gga	agt	gga	1064
119	Glu	Asn	Cys	Thr	Phe	Gly	Gly	Ile	Trp	Asp	Gly	Gly	Lys	Gly	Ser	Gly	
120	275				280					285					290		
122	cag	aaa	aac	ctt	gtt	act	tca	gct	ttc	tac	tat	agg	tct	tct	gag		1112
123	Gln	Lys	Asn	Leu	Val	Val	Thr	Ser	Ala	Phe	Tyr	Tyr	Arg	Ser	Ser	Glu	
124					295					300					305		
126	gtt	ggt	ttt	gtc	act	cct	ccc	aat	tcc	aaa	aat	cgc	cct	ctg	gat	ttt	1160

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/657,631

DATE: 03/13/2002
TIME: 14:59:12

Input Set : A:\-798-2.app
Output Set: N:\CRF3\03132002\I657631.raw

127 Val Gly Phe Val Thr Pro Pro Asn Ser Lys Asn Arg Pro Leu Asp Phe
128 310 315 320
130 gaa act gca gct aaa caa gct tgt agt tta aca ttc gag gaa gcg aaa 1208
131 Glu Thr Ala Ala Lys Gln Ala Cys Ser Leu Thr Phe Glu Glu Ala Lys
132 325 330 335
134 tcc act ttt cca aat gtt gag aaa gat aaa ctt cca ttt gta tgc gtg 1256
135 Ser Thr Phe Pro Asn Val Glu Lys Asp Lys Leu Pro Phe Val Cys Val
136 340 345 350
138 gat ttc aca tac cag tat aca ttg ctt gtt gat gga ttt ggc cta gat 1304
139 Asp Phe Thr Tyr Gln Tyr Thr Leu Leu Val Asp Gly Phe Gly Leu Asp
140 355 360 365 370
142 cca gag caa gag att aca gtg gca gaa gga att gaa tat caa gat gcc 1352
143 Pro Glu Gln Glu Ile Thr Val Ala Glu Gly Ile Glu Tyr Gln Asp Ala
144 375 380 385
146 att gtg gaa aca gca tgg cct cta gga act gcc ata gaa gcc ata tca 1400
147 Ile Val Glu Thr Ala Trp Pro Leu Gly Thr Ala Ile Glu Ala Ile Ser
148 390 395 400
150 tct ttg cct aaa ttt aat cgt cta atg tat ttt atc taa gccatgtcct 1449
151 Ser Leu Pro Lys Phe Asn Arg Leu Met Tyr Phe Ile
W--> 152 405 410 415
154 ccacttatga ccacttaat taaaataaaa ctcacccttt tcactaaaaaa aaaaaaaaaa 1509
156 aaaagtccctt ttttattcca ttgagtatca agtgttaatt tgggtctgac aaatggaggt 1569
158 gtaaaaagtga aacaaagtat gttttgtca gatacgaatg gaagtagggg tatgatgaaa 1629
160 aaaaaaaaaaaa aaaa 1643
163 <210> SEQ ID NO: 2
164 <211> LENGTH: 462
165 <212> TYPE: PRT
166 <213> ORGANISM: Dolichos biflorus
168 <220> FEATURE:
169 <223> OTHER INFORMATION: lectin nucleotide phosphohydrolase (LNP, NBP46 or
170 DB46) root lectin
172 <400> SEQUENCE: 2
173 Met Asn Trp Val Trp Pro Lys Thr Lys Ser Met Ser Phe Leu Leu Leu
174 1 5 10 15
175 Ile Thr Phe Leu Leu Phe Ser Leu Pro Lys Leu Ser Ser Ser Gln Tyr
176 20 25 30
177 Val Gly Asn Ser Ile Leu Leu Asn His Arg Lys Ile Leu Pro Asn Gln
178 35 40 45
179 Glu Leu Leu Thr Ser Tyr Ala Val Ile Phe Asp Ala Gly Ser Ser Gly
180 50 55 60
181 Ser Arg Val His Val Phe Asn Phe Asp Gln Asn Leu Asp Leu Leu His
182 65 70 75 80
183 Ile Gly Asn Asp Leu Glu Phe Thr Lys Lys Ile Lys Pro Gly Leu Ser
184 85 90 95
185 Ser Tyr Ala Asp Lys Pro Glu Lys Ala Ala Glu Ser Leu Ile Pro Leu
186 100 105 110
187 Leu Glu Glu Ala Glu Asp Val Val Pro Glu Glu Leu His Pro Lys Thr
188 115 120 125
189 Pro Leu Lys Leu Gly Ala Thr Ala Gly Leu Arg Leu Leu Asp Gly Asp

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/657,631

DATE: 03/13/2002
TIME: 14:59:12

Input Set : A:\-798-2.app
Output Set: N:\CRF3\03132002\I657631.raw

190 130 135 140
 191 Ala Ala Glu Lys Ile Leu Gln Ala Val Arg Glu Met Phe Arg Asn Arg
 192 145 150 155 160
 193 Ser Ser Leu Ser Val Gln Pro Asp Ala Val Ser Val Ile Asp Gly Thr
 194 165 170 175
 195 Gln Glu Gly Ser Tyr Leu Trp Val Thr Val Asn Tyr Leu Leu Gly Lys
 196 180 185 190
 197 Leu Gly Lys Phe Thr Lys Thr Val Gly Val Ile Asp Leu Gly Gly
 198 195 200 205
 199 Ala Ser Val Gln Met Ala Tyr Ala Val Ser Arg Asn Thr Ala Lys Asn
 200 210 215 220
 201 Ala Pro Lys Pro Pro Gln Gly Glu Asp Pro Tyr Met Lys Lys Leu Val
 202 225 230 235 240
 203 Leu Lys Gly Lys Lys Tyr Asp Leu Tyr Val His Ser Tyr Leu Arg Tyr
 204 245 250 255
 205 Gly Asn Asp Ala Ala Arg Val Lys Ile Phe Lys Thr Thr Asp Gly Ala
 206 260 265 270
 207 Ala Ser Pro Cys Leu Leu Ala Gly Tyr Glu Asp Ile Tyr Arg Tyr Ser
 208 275 280 285
 209 Gly Glu Ser Tyr Asn Ile Tyr Gly Pro Thr Ser Gly Ala Asn Phe Asn
 210 290 295 300
 211 Glu Cys Arg Asp Leu Ala Leu Gln Ile Leu Arg Leu Asn Glu Pro Cys
 212 305 310 315 320
 213 Ser His Glu Asn Cys Thr Phe Gly Gly Ile Trp Asp Gly Gly Lys Gly
 214 325 330 335
 215 Ser Gly Gln Lys Asn Leu Val Val Thr Ser Ala Phe Tyr Tyr Arg Ser
 216 340 345 350
 217 Ser Glu Val Gly Phe Val Thr Pro Pro Asn Ser Lys Asn Arg Pro Leu
 218 355 360 365
 219 Asp Phe Glu Thr Ala Ala Lys Gln Ala Cys Ser Leu Thr Phe Glu Glu
 220 370 375 380
 221 Ala Lys Ser Thr Phe Pro Asn Val Glu Lys Asp Lys Leu Pro Phe Val
 222 385 390 395 400
 223 Cys Val Asp Phe Thr Tyr Gln Tyr Thr Leu Leu Val Asp Gly Phe Gly
 224 405 410 415
 225 Leu Asp Pro Glu Gln Glu Ile Thr Val Ala Glu Gly Ile Glu Tyr Gln
 226 420 425 430
 227 Asp Ala Ile Val Glu Thr Ala Trp Pro Leu Gly Thr Ala Ile Glu Ala
 228 435 440 445
 229 Ile Ser Ser Leu Pro Lys Phe Asn Arg Leu Met Tyr Phe Ile
 230 450 455 460
 233 <210> SEQ ID NO: 3
 234 <211> LENGTH: 1458
 235 <212> TYPE: DNA
 236 <213> ORGANISM: Medicago sativa
 238 <220> FEATURE:
 239 <221> NAME/KEY: CDS
 240 <222> LOCATION: (1)..(1458)
 241 <223> OTHER INFORMATION: full length clone

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/657,631

DATE: 03/13/2002
TIME: 14:59:12

Input Set : A:\-798-2.app
Output Set: N:\CRF3\03132002\I657631.raw

243 <220> FEATURE:
244 <221> NAME/KEY: CDS
245 <222> LOCATION: (13)..(1380)
246 <223> OTHER INFORMATION: lectin nucleotide phosphohydrolase (LNP or NBP46)
247 root lectin
249 <400> SEQUENCE: 3
250 caa att aag aac atg gag ttc cta att aca ctc att gcc act ttt tta 48
251 Gln Ile Lys Asn Met Glu Phe Leu Ile Thr Leu Ile Ala Thr Phe Leu
252 1 5 10 15
254 ctc ttg tta atg cct gca atc act tcc tcc caa tat tta gga aac aac 96
255 Leu Leu Leu Met Pro Ala Ile Thr Ser Ser Gln Tyr Leu Gly Asn Asn
256 20 25 30
258 cta ctc act aat cga aag att ttc caa aaa caa gaa acc tta acc tct 144
259 Leu Leu Thr Asn Arg Lys Ile Phe Gln Lys Gln Glu Thr Leu Thr Ser
260 35 40 45
262 tac gct gtc ata ttt gat gct ggt agc act ggt act cgt gtc cat gtt 192
263 Tyr Ala Val Ile Phe Asp Ala Gly Ser Thr Gly Thr Arg Val His Val
264 50 55 60
266 tac cat ttt gat cag aac tta gat cta ctt cac att ggc aat gat att 240
267 Tyr His Phe Asp Gln Asn Leu Asp Leu Leu His Ile Gly Asn Asp Ile
268 65 70 75 80
270 gag ttt gtt gac aag atc aaa cca ggt ttg agt gca tat ggg gat aat 288
271 Glu Phe Val Asp Lys Ile Lys Pro Gly Leu Ser Ala Tyr Gly Asp Asn
272 85 90 95
274 cct gaa caa gca gca aaa tct ctc att cca ctt ttg gag gaa gca gaa 336
275 Pro Glu Gln Ala Ala Lys Ser Leu Ile Pro Leu Leu Glu Ala Glu
276 100 105 110
278 gat gtg gtt cct gag gat ctg cac ccc aaa aca ccc ctt agg ctt ggg 384
279 Asp Val Val Pro Glu Asp Leu His Pro Lys Thr Pro Leu Arg Leu Gly
280 115 120 125
282 gca acc gca ggt ttg agg ctt ttg aat ggg gat gct gct gaa aag ata 432
283 Ala Thr Ala Gly Leu Arg Leu Leu Asn Gly Asp Ala Ala Glu Lys Ile
284 130 135 140
286 ttg caa gcg aca agg aat atg ttc agc aac aga agt acc ctc aac gtt 480
287 Leu Gln Ala Thr Arg Asn Met Phe Ser Asn Arg Ser Thr Leu Asn Val
288 145 150 155 160
290 caa cgt gat gca gtt tct att att gat gga acc caa gaa ggt tct tat 528
291 Gln Arg Asp Ala Val Ser Ile Ile Asp Gly Thr Gln Glu Gly Ser Tyr
292 165 170 175
294 atg tgg gtg aca gtt aac tat gta ttg ggg aat ttg gga aaa agc ttc 576
295 Met Trp Val Thr Val Asn Tyr Val Leu Gly Asn Leu Gly Lys Ser Phe
296 180 185 190
298 aca aaa tca gtg gga gta att gac ctt gga ggt tca gtt caa atg 624
299 Thr Lys Ser Val Gly Val Ile Asp Leu Gly Gly Ser Val Gln Met
300 195 200 205
302 aca tat gca gtg tca aag aaa aca gca aaa aat gct cct aaa gtt gct 672
303 Thr Tyr Ala Val Ser Lys Lys Thr Ala Lys Asn Ala Pro Lys Val Ala
304 210 215 220
306 gat gga gag gat cca tat att aag aag ctt gtg ctc aag gga aag caa 720

Use of n and / or Xaa has been detected in the
Sequence Listing. Review the Sequence Listing
to ensure a corresponding explanation is present
in the <220> to <223> fields of each sequence
using n or Xaa.

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/657,631

DATE: 03/13/2002
TIME: 14:59:13

Input Set : A:\-798-2.app
Output Set: N:\CRF3\03132002\I657631.raw

L:152 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:364 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:368 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:372 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:494 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:498 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:502 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:506 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:510 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:514 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:518 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:522 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:526 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:530 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:534 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:538 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:542 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:546 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:550 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:554 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:558 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:562 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:566 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:570 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:574 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:578 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:582 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:586 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:590 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:594 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:598 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:602 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:606 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:610 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:614 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8
L:737 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:755 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14